

Bayes' Rule

CSE 312 Summer 25
Lecture 6

Announcements

Optional Quiz is available on Ed after lecture.

Quiz 2 will be on Friday.

Homework 1 solutions available after lecture.

Last Time

Conditional Probability

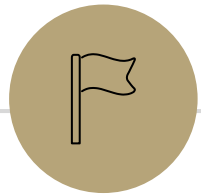
For an event B , with $\mathbb{P}(B) > 0$,
the “Probability of A conditioned on B ” is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Roll a fair 6-sided die. Let A be “the die roll is a 4”, B be “the die roll is an even number.”

$$P(A|B) =$$

$$P(B|A) =$$



Bayes' Rule



Wonka Bars

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

You want to get a golden ticket. You could buy a 1000-or-so of the bars until you find one, but that's expensive...you've got a better idea!

You have a test – a very precise scale you've bought.

- If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.
- If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

If you pick up a bar and it alerts, what is the probability you have a golden ticket?

Willy Wonka

When you have an idea
go to pollev.com/avk5

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

- If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.
- If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

You pick up a bar and it alerts, what is the probability you have a golden ticket?

Which of these is closest to the right answer?

- A. 0.1%
- B. 10%
- C. 50%
- D. 90%
- E. 99%
- F. 99.9%

Conditioning

Let S be the event that the Scale alerts you

Let G be the event your bar has a Golden ticket.

What conditional probabilities are each of these?

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

$$\mathbb{P}(G)$$

If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.

$$\mathbb{P}(S|G)$$

If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

$$\mathbb{P}(S|\bar{G})$$

You pick up a bar and it alerts, what is the probability you have a golden ticket?

$$\mathbb{P}(G|S)$$

Reversing the Conditioning

All of our information conditions on whether G happens or not – does your bar have a golden ticket or not?

But we're interested in the "reverse" conditioning. We know the scale alerted us – we know the test is positive – but do we have a golden ticket?

Bayes' Rule

Bayes' Rule

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

Bayes' Rule

Bayes' Rule

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{\mathbb{P}(S|G) \cdot \mathbb{P}(G)}{\mathbb{P}(S)}$$

Bayes' Rule

Bayes' Rule

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$$

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{.999 \cdot .001}{\mathbb{P}(S)}$$

Filling In

What's $\mathbb{P}(S)$?

We'll use a trick called "the law of total probability":

$$\begin{aligned}\mathbb{P}(S) &= \mathbb{P}(S|G) \cdot \mathbb{P}(G) + \mathbb{P}(S|\bar{G}) \cdot \mathbb{P}(\bar{G}) \\ &= 0.999 \cdot .001 + .01 \cdot .999 \\ &= .010989\end{aligned}$$

Law of Total Probability

Let A_1, A_2, \dots, A_k be a partition of Ω .

A partition of a set S is a family of subsets S_1, S_2, \dots, S_k such that:

$S_i \cap S_j = \emptyset$ for all i, j and

$S_1 \cup S_2 \cup \dots \cup S_k = S$.

i.e. every element of Ω is in exactly one of the A_i .

Law of Total Probability

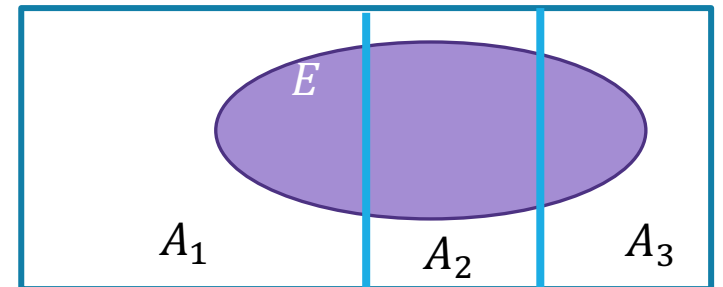
Law of Total Probability

Let A_1, A_2, \dots, A_k be a partition of Ω .

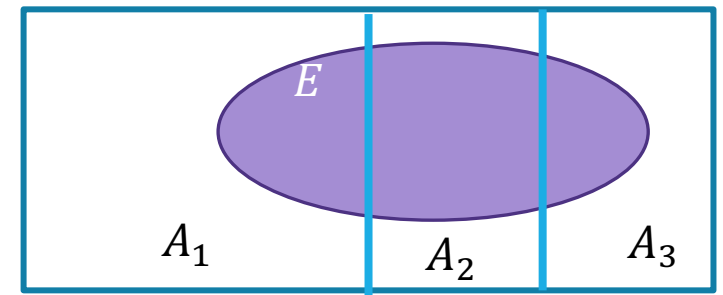
For any event E ,

$$\mathbb{P}(E) = \sum_{\text{all } i} \mathbb{P}(E|A_i)\mathbb{P}(A_i)$$

Ω , split into partition A_1, A_2, A_3 with event E inside.



Why?



The Proof is actually pretty informative on what's going on.

$$\begin{aligned} & \sum_{\text{all } i} \mathbb{P}(E|A_i)\mathbb{P}(A_i) \\ &= \sum_{\text{all } i} \frac{\mathbb{P}(E \cap A_i)}{\mathbb{P}(A_i)} \cdot \mathbb{P}(A_i) \text{ (definition of conditional probability)} \\ &= \sum_{\text{all } i} \mathbb{P}(E \cap A_i) \\ &= \mathbb{P}(E) \end{aligned}$$

The A_i partition Ω , so $E \cap A_i$ partition E . Then we just add up those probabilities.

Ability to add follows from the "countable additivity" axiom.

Bayes Rule

What do we know about Wonka Bars?

$$\mathbb{P}(G|S) = \frac{.999 \cdot .001}{.010989}$$

Solving $\mathbb{P}(G|S) = \frac{1}{11}$, i.e. about 0.0909.

Only about a 10% chance that the bar has the golden ticket!

Wait a minute...

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars. If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time. If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

That doesn't fit with many of our guesses. What's going on?

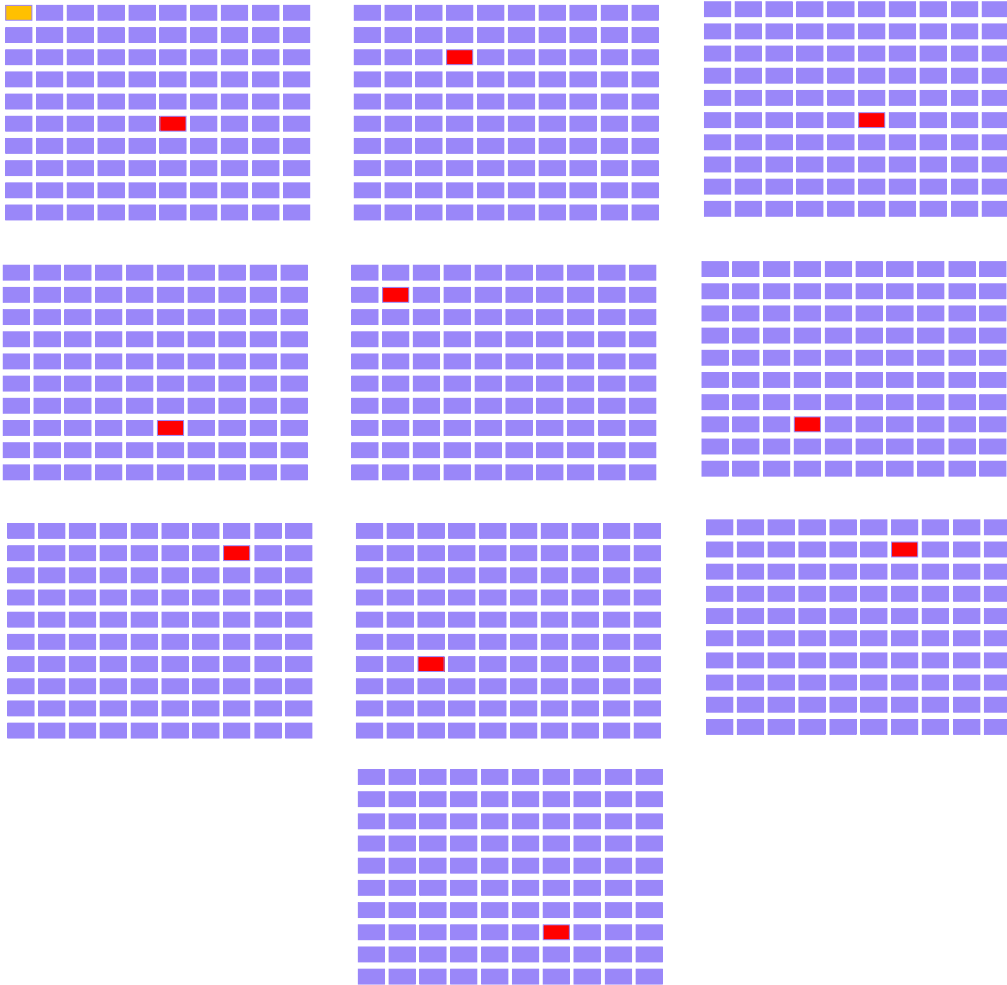
Instead of saying "we tested one and got a positive" imagine we tested 1000. **ABOUT** how many bars of each type are there?

(about) 1 with a golden ticket 999 without. Let's say those are exactly right.

Let's just say that one golden is truly found.

(about) 1% of the 999 without would be a positive. Let's say it's exactly 10.

Visually



Gold bar is the one (true) golden ticket bar. Purple bars don't have a ticket and tested negative.

Red bars don't have a ticket, but tested positive.

The test is, in a sense, doing really well. It's almost always right.

The problem is it's also the case that the correct answer is almost always "no."

Updating Your Intuition

🔥 Take 1: The test is **actually good** and has VASTLY increased our belief that there IS a golden ticket when you get a positive result.

If we told you “your job is to find a Wonka Bar with a golden ticket” without the test, you have 1/1000 chance, with the test, you have (about) a 1/11 chance. That’s (almost) 100 times better!

This is actually a huge improvement!

Updating Your Intuition

🔥 Take 2: Humans are really bad at intuitively understanding very large or very small numbers.

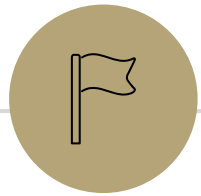
When I hear "99% chance", "99.9% chance", "99.99% chance" they all go into my brain as "well that's basically guaranteed" And then I forget how many 9's there actually were.

Updating Your Intuition

 Take 3: View tests as updating your beliefs, not as revealing the truth.

Bayes' Rule says that $\mathbb{P}(B|A)$ has a factor of $\mathbb{P}(B)$ in it. You have to translate "The test says there's a golden ticket" to "the test says you should increase your estimate of the chances that you have a golden ticket."

A test takes you from your "prior" beliefs of the probability to your "posterior" beliefs.



More Bayes Practice

A contrived example

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket.

You will flip a fair coin; if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.

Let B be you draw a blue marble. Let T be the coin is tails.

What is $\mathbb{P}(B|T)$ what is $\mathbb{P}(T|B)$?

Flipping the conditioning

What about $\mathbb{P}(T|B)$?

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket. if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.

Pause, what's your intuition?

Is this probability

A. less than $\frac{1}{2}$

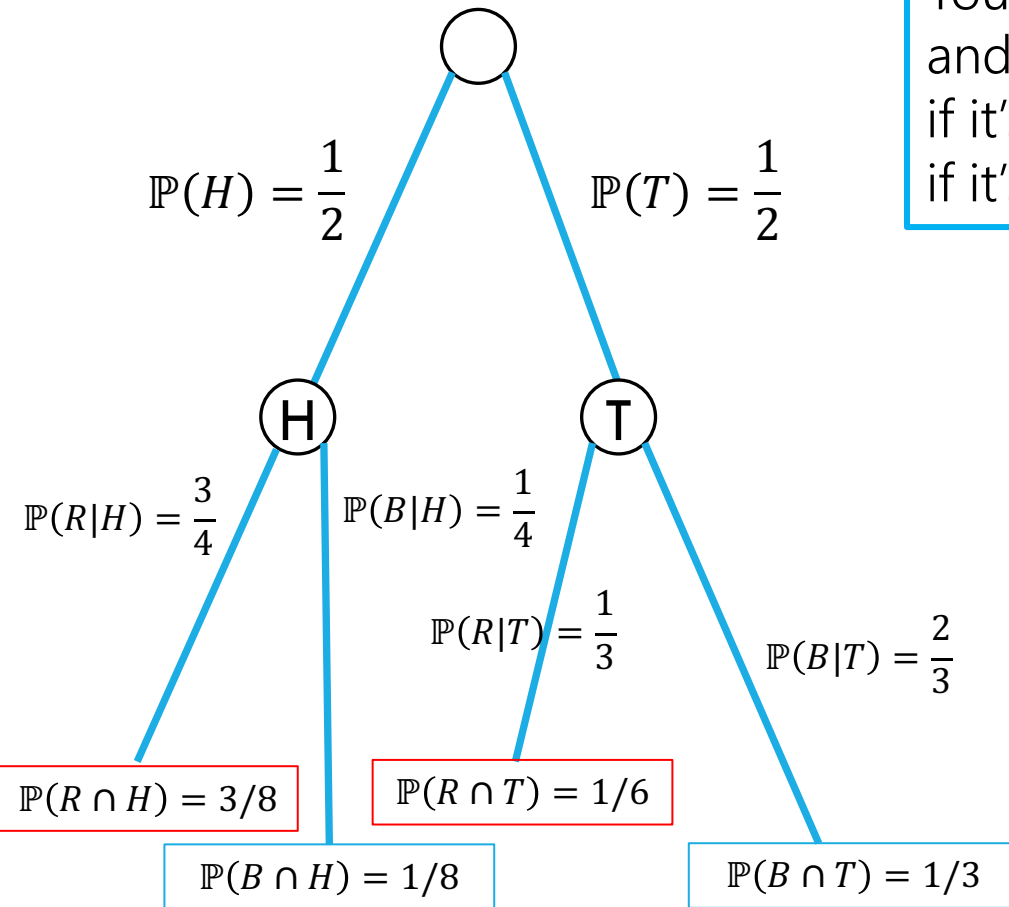
B. equal to $\frac{1}{2}$

C. greater than $\frac{1}{2}$

The right (tails) pocket is far more likely to produce a blue marble if picked than the left (heads) pocket is. Seems like $\mathbb{P}(T|B)$ should be greater than $\frac{1}{2}$.

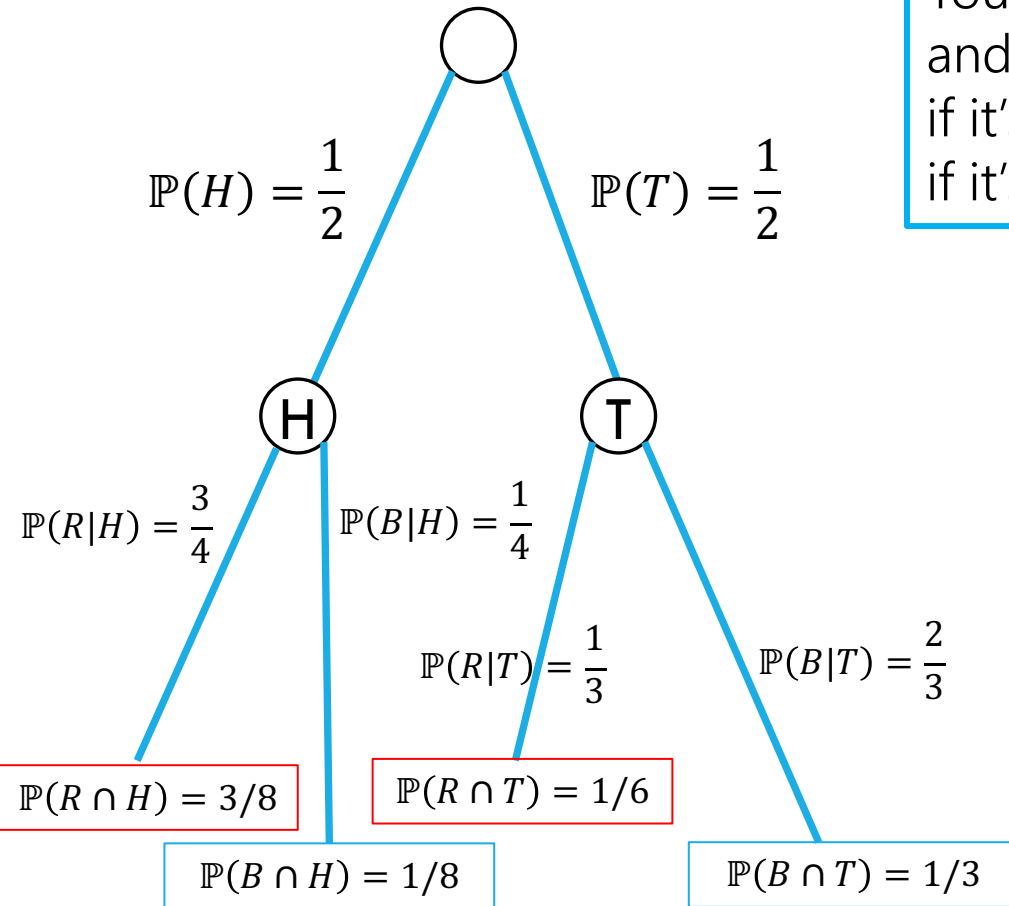
Updated Sequential Processes

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket. if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.



Updated Sequential Processes

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket. if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.



$$\mathbb{P}(B|T) = \frac{2}{3}$$

$$\mathbb{P}(B) = \frac{1}{8} + \frac{1}{3} = \frac{11}{24}$$

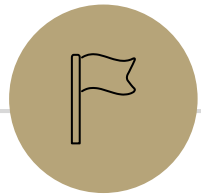
Flipping the conditioning

What about $\mathbb{P}(T|B)$?

You have three red marbles and one blue marble in your left pocket, and one red marble and two blue marbles in your right pocket. if it's heads, you'll draw a marble (uniformly) from your left pocket, if it's tails, you'll draw a marble (uniformly) from your right pocket.

Bayes' Rule says:

$$\begin{aligned}\mathbb{P}(T|B) &= \frac{\mathbb{P}(B|T)\mathbb{P}(T)}{\mathbb{P}(B)} \\ &= \frac{\frac{2}{3} \cdot \frac{1}{2}}{11/24} \\ &= 8/11\end{aligned}$$



The Technical Stuff

Proof of Bayes' Rule

$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ by definition of conditional probability

Now, imagining we get $A \cap B$ by conditioning on A , we should get a numerator of $\mathbb{P}(B|A) \cdot \mathbb{P}(A)$

$$= \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

As required.

A Technical Note

After you condition on an event, what remains is a probability space.

With B playing the role of the sample space,

$\mathbb{P}(\omega|B)$ playing the role of the probability measure.

All the axioms are satisfied (it's a good exercise to check)

That means any theorem we write down has a version where you condition everything on B .

An Example

Bayes Theorem still works in a probability space where we've already conditioned on S .

$$\mathbb{P}(A|[B \cap S]) = \frac{\mathbb{P}(B|[A \cap S]) \cdot \mathbb{P}(A|S)}{\mathbb{P}(B|S)}$$

Complementary law still works in a probability space where we've already conditioned on S

$$\mathbb{P}(A|C) = 1 - \mathbb{P}(\bar{A}|C)$$

A Quick Technical Remark

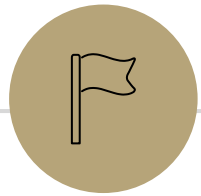
Sometimes students write things like

$$\mathbb{P}([A|B]|C)$$

This is not a thing.

You probably want $\mathbb{P}(A|[B \cap C])$

$A|B$ isn't an event – it's describing an event **and** telling you to restrict the sample space. So you can't ask for the probability of that conditioned on something else.



Bayes in the real world



Application 1: Medical Tests

Helping Doctors and Patients Make Sense of Health Statistics

A researcher posed the following scenario to a group of 160 doctors:

Assume you conduct a disease screening using a standard test in a certain region. You know the following information about the people in this region:

The probability that a person has the disease is 1% (prevalence)

If a person has the disease, the probability that she tests positive is 90% (sensitivity)

If a person does not have the disease, the probability that she nevertheless tests positive is 9% (false-positive rate)

A person tests positive. She wants to know from you whether that means that she has the disease for sure, or what the chances are. What is the best answer?

A. The probability that she has the disease is about 81%.

B. Out of 10 people with a positive test, about 9 have the disease.

C. Out of 10 people with a positive test, about 1 have the disease.

D. The probability that she has the disease is about 1%

Let's do the calculation!

Let D be "the patient has the disease", T be the test was positive.

$$\begin{aligned}\mathbb{P}(D|T) &= \mathbb{P}(T|D) \cdot \mathbb{P}(D) / \mathbb{P}(T) \\ &= \frac{.9 \cdot .01}{.99 \cdot .09 + .01 \cdot .9} \approx 0.092\end{aligned}$$

Calculation tip: for Bayes' Rule, you should see one of the terms on the bottom exactly match your numerator (if you're using the LTP to calculate the probability on the bottom)

Pause for vocabulary

Physicians have words for just about everything

Let D be has the disease; T be test is positive

$\mathbb{P}(D)$ is "prevalence"

$\mathbb{P}(T|D)$ is "sensitivity"

A 'sensitive' test is one which picks up on the disease when it's there
(high sensitivity -> few false negatives)

$\mathbb{P}(\bar{T}|\bar{D})$ is "specificity"

A 'specific' test is one that is positive specifically because of the disease, and for no other reason (high specificity -> few false positives)

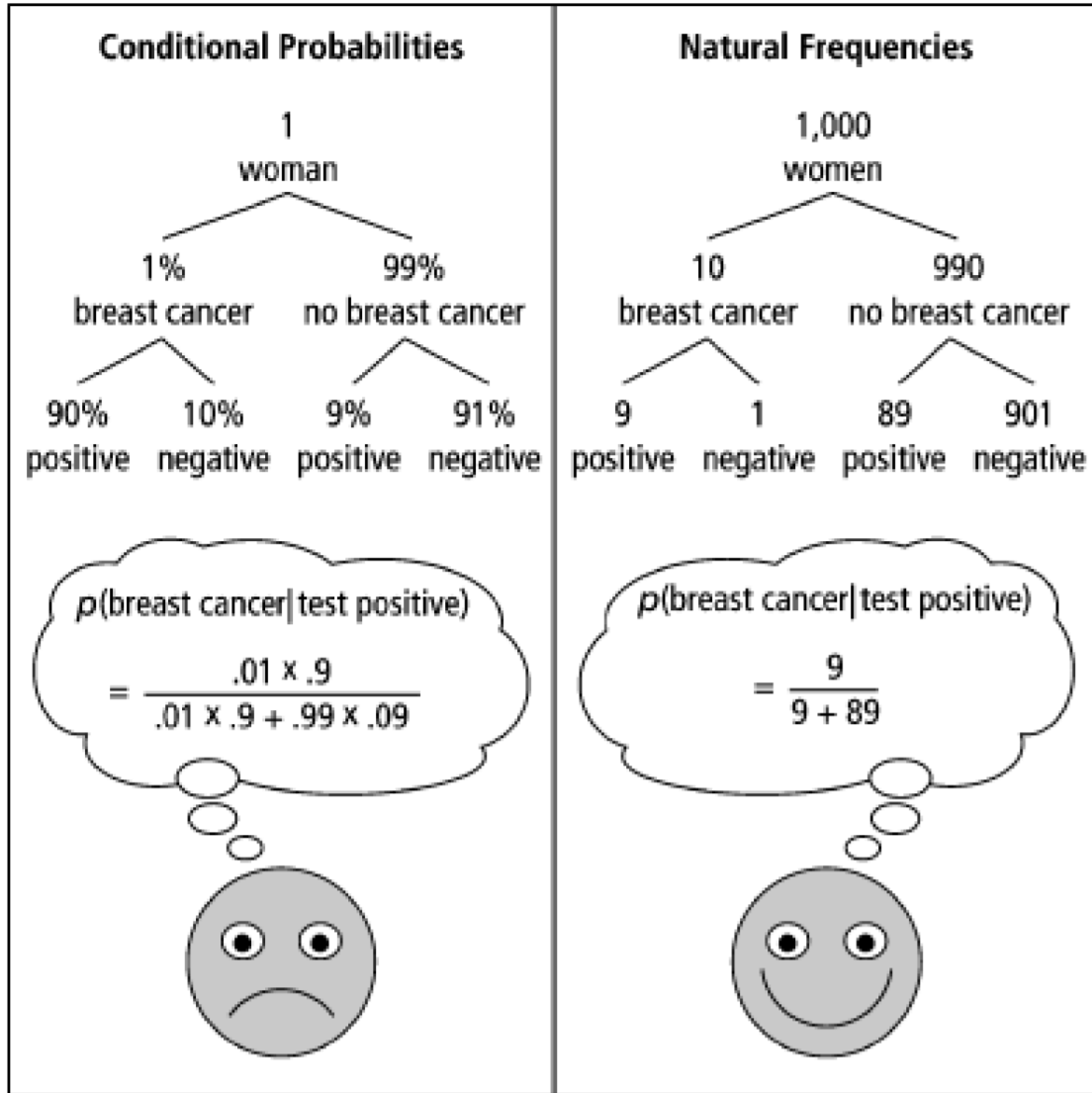
How did the doctors do

C (about 1 in 10) was the correct answer.

Of the doctors surveyed, less than $\frac{1}{4}$ got it right (so worse than random guessing).

After the researcher taught them his calculation trick, more than 80% got it right.

One Weird Trick!



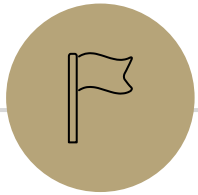
Calculation Trick: imagine you have a large population (not one person) and ask how many there are of false/true positives/negatives.

What about the real world?

When you're older and have to do more routine medical tests, don't get concerned (yet) when they ask to run another test.*

It's usually fine.*

*This is not medical advice, Anna is not a physician.



Optional: Bayes Factor

A way to estimate Bayes calculations quickly

Bayes Factor

Another Intuition Trick: [from 3Blue1Brown](#)

When you test positive, you (**approximately**) multiply the prior by the “Bayes Factor” (aka likelihood ratio)

$$\frac{\text{sensitivity}}{\text{false positive rate}} = \frac{1 - FNR}{FPR}$$

Bayes Factor

Does it work?

Let's try it...

Find

$$\text{prior} \cdot \frac{\text{Sensitivity}}{FPR}$$

Wonka Bars

Willy Wonka has placed golden tickets on 0.1% of his Wonka Bars.

You want to get a golden ticket. You could buy a 1000-or-so of the bars until you find one, but that's expensive...you've got a better idea!

You have a test – a very precise scale you've bought.

If the bar you weigh **does** have a golden ticket, the scale will alert you 99.9% of the time.

If the bar you weigh does not have a golden ticket, the scale will (falsely) alert you only 1% of the time.

If you pick up a bar and it alerts, what is the probability you have a golden ticket?

Wonka Bars

Bayes Factor

$$\frac{99.9}{1}$$

Prior: .1%

Product: 9.99, so about 10%

About what Bayes Rule gets!

Application 1: Medical Tests

Helping Doctors and Patients Make Sense of Health Statistics

A researcher posed the following scenario to a group of 160 doctors:

Assume you conduct a disease screening using a standard test in a certain region. You know the following information about the people in this region:

The probability that a person has the disease is 1% (prevalence)

If a person has the disease, the probability that she tests positive is 90% (sensitivity)

If a person does not have the disease, the probability that she nevertheless tests positive is 9% (false-positive rate)

A person tests positive. She wants to know from you whether that means that she has the disease for sure, or what the chances are. What is the best answer?

A. The probability that she has the disease is about 81%.

B. Out of 10 people with a positive test, about 9 have the disease.

C. Out of 10 people with a positive test, about 1 have the disease.

D. The probability that she has the disease is about 1%

Bayes Factor

What about with the doctors?

$$1\% \cdot \frac{90\%}{9\%} = 10\%$$

Again about right!

Caution

Multiplying by the Bayes Factor is an **approximation**

It gives you the exact numerator for Bayes, but the denominator is "the number of false positives if the prevalence (/prior) were 0"

When the prior is close to 0, this is a fine approximation!

But plug in a prior of 15% on the last slide, and we get 150% chance.

What about negative tests?

For negative tests, the Bayes Factor is

$$\frac{FNR}{\text{specificity}}$$

Specificity is (1 – false negative rate)